*NOTE* **Communicated by Anthony Bell**

# The Shape of Neural Dependence

**Rick L. Jenison**
*rjenison@wisc.edu*
**Richard A. Reale**
*reale@physiology.wisc.edu*
*Departments of Psychology and Physiology and the Waisman Center, University of Wisconsin–Madison, Madison, WI 53706, U.S.A.*

**The product-moment correlation coefficient is often viewed as a natural measure of dependence. However, this equivalence applies only in the context of elliptical distributions, most commonly the multivariate gaussian, where linear correlation indeed sufficiently describes the underlying dependence structure. Should the true probability distributions deviate from those with elliptical contours, linear correlation may convey misleading information on the actual underlying dependencies. It is often the case that probability distributions other than the gaussian distribution are necessary to properly capture the stochastic nature of single neurons, which as a consequence greatly complicates the construction of a flexible model of covariance. We show how arbitrary probability densities can be coupled to allow greater flexibility in the construction of multivariate neural population models.**

---

Neural population coding of events in the world is thought to involve a coordinated activity among neurons. In delineating the structure of this dependence among neurons in the brain, it is the time pattern of discharges from individual cells that represents the variables or marginal dimensions under study. A common misperception is that any marginal probability distribution can be used to construct a parametric model of this dependence among variables. The most commonly used measure for dependence, the Pearson product-moment correlation coefficient, was developed on the basis of normal marginals and addresses only linear dependence (Yule, 1897; Mari & Kotz, 2001). However, a new and innovative approach, the so-called copula method, provides the ability to couple arbitrary marginal densities (Sklar, 1959; Joe, 1997; Nelsen, 1999). The word *copula* is a Latin noun that refers to a bond and is used in linguistics to refer to a proposition that links a subject and predicate. In probability theory, it couples marginal distributions to form flexible multivariate distribution functions. The appeal of the copula is that we eliminate the implied reliance on the multivariate gaussian or the assumption that dimensions are independent.

 

Recently, Nirenberg, Carcieri, Jacobs, and Latham (2001) indirectly estimated the relative contribution of neural dependence to mutual information, which accounted for about 10% of the information transmitted. Pola, Thiele, Hoffmann, and Panzeri (2003) derived a method to directly estimate the mutual information on the entire response joint probability distribution. In both of these reports, as well as earlier studies of neural information, the shape of the dependence structure is generally ignored and considered to be absorbed into the joint probability density. In this communication, we show how the conditional response entropy can be factored into separate independent and dependent contributions to transmitted sensory information. These factors take the form of conditional differential entropies, and therefore the structure and impact of the dependence can be computed directly rather than indirectly.

Sklar's theorem (Sklar, 1959) states that any multivariate distribution can be expressed as the copula function $\mathbf{C}[u_1, u_2, \ldots, u_N]$ evaluated at each of the marginal distributions. By virtue of the probability integral transform (Casella & Berger, 1990), each marginal $u_i = F_i(x_i)$ has a uniform distribution on $[0, 1]$, where $F_i(x_i)$ is the cumulative integral of $p_i[x_i]$ for the random variables $X_i$. The joint probability density follows as

$$p[x_1, x_2, \ldots, x_N] = \prod_{i=1}^{N} p_i[x_i] \times \mathrm{c}[u_1, u_2, \ldots, u_N], \tag{1}$$

where $p_i[x_i]$ is each marginal density and coupling is provided by $\mathrm{c}[u_1, u_2, \ldots, u_N] = \frac{\partial^N \mathbf{C}[u_1, u_2, \ldots, u_N]}{\partial u_1 \partial u_2 \cdots \partial u_N}$, which is itself a probability density. When the random variables are independent, the copula density $\mathrm{c}[u_1, u_2, \ldots, u_N]$ is identically equal to one. The importance of equation 1 is that the independent portion, reflected as the product of the marginals, can be separated from the function $\mathrm{c}[u_1, u_2, \ldots, u_N]$ describing the dependence structure or shape. There are several families of copulas, some of which are particularly descriptive of the dependence structure observed among the discharge patterns of sensory neurons. The choice of copula is a form of model selection that can be formalized using information criteria (Soofi, 2000) such as the Akaike information criterion (AIC) (Akaike, 1974). The AIC penalizes the negative log maximum likelihood of the estimated model by the number of parameters in the model [−2 log (maximum likelihood) + 2 (number of parameters)]. A smaller relative AIC represents a better model fit while taking into account the complexity of the model.

We have shown previously the importance of considering the impact of correlated noise on the acuity of sound localization by an ensemble of auditory cortical neurons in the cat (Jenison, 2000). Figure 1A shows a scatter plot of the timing (latency) to the first evoked spike following a sound presented at a particular location, $\theta$, in space for a pair of neurons recorded from separate electrodes in the primary auditory (AI) field of an anesthetized cat.

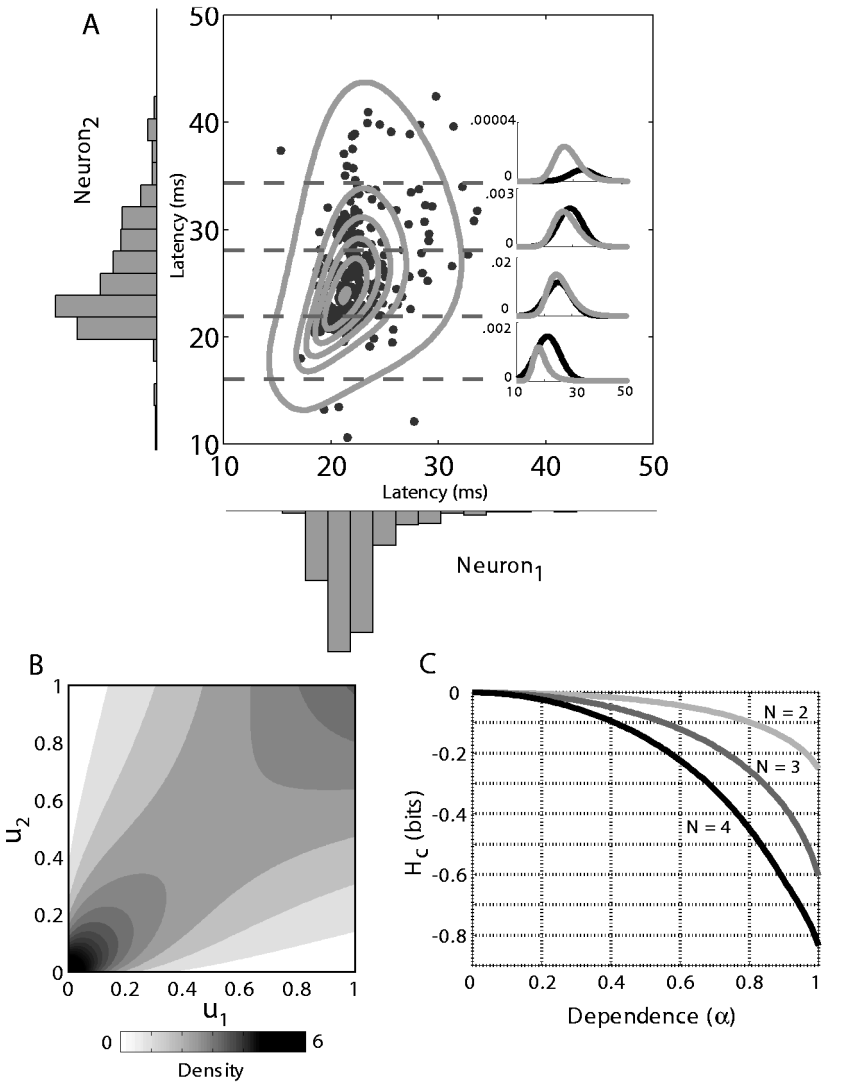


Figure 1: (A) Joint conditional probability density of first-spike latency recorded from two cortical neurons in response to a given sound source in space. Gray contours correspond to estimated iso-densities based on the AMH copula. The estimated single parameter $\alpha$ controls the shape of the copula, here $\alpha = 0.95$. The inset shows conditional cross sections of the joint density based on the copula (gray) and estimated multivariate gaussian (black) (B) AMH copula density $c[u_1, u_2 \mid \theta]$ estimated from data shown in $A$. (C) AMH copula conditional entropy as a function of dependence and neural ensemble size ($N$). Multiple integrals were evaluated numerically using quasi–Monte Carlo integration.

These two neurons are representative of a common class of field AI cells described by their spatial sensitivity to free-field (actual or virtual) sound sources. In this regard, these AI neurons exhibited broad spatial receptive fields when measured with directionally dependent stimuli even at the low intensity levels of the sound source (Brugge, Reale, & Hind, 1996). Furthermore, their response latency was tightly locked to stimulus onset, and systematic gradients in latency were evident within the receptive field. Models have been successfully developed that capture this systematic response latency variability due to both sound-source direction and sound-source intensity (Jenison, Reale, Hind, & Brugge, 1998; Reale, Jenison, & Brugge, 2003).

The joint probability density contours (gray) reflect maximum likelihood fit inverse-gaussian (IG) marginal probability densities (Jenison & Reale, 2003) coupled by the Ali-Mikail-Haq (AMH) copula (see appendix A). The AMH copula is parameterized by a dependence measure $\alpha$, which for the general multivariate case ranges between 0 and 1. The fit shown in Figure 1 yields $\alpha = 0.95$ with an AIC equal to 3642. In comparison, for this data set, the AIC for independent IG marginals is 3728, and the multivariate gaussian AIC is 3766. This figure shows two common characteristics of the joint response of auditory cortical neurons. First, the marginal distributions are positively skewed, unlike the gaussian distribution, and second, the dependence structure is not elliptical, which is also inconsistent with a multivariate gaussian density. The inset shows comparisons of this copula-based joint density and the corresponding multivariate gaussian fit conditioned on several levels of first-spike latencies from neuron 2. The estimated AMH copula density $c[u_1, u_2]$ is shown in Figure 1B, which can be viewed as a function that modulates the product density to form the joint probability density function as defined by equation 1. In this case, the AMH copula density narrows in the lower tails (near zero) and broadens in the upper tails (near one). This is a characteristic of the AMH copula and relates to our general observations of AI cortical neurons where the strength of association between ensemble neurons is greater for shorter first-spike latencies compared to responses of longer latencies. When a pair of neurons fires in response to a particular location in space, the shorter latencies (which reflect a strong response) are more consistent within the ensemble; longer latencies (weaker responses) tend to be less consistent. This analysis indicates that not only is a multivariate gaussian density no longer a necessary assumption in delineating the structure of dependence among neurons, but employing a multivariate gaussian density with these data would actually provide misleading information, particularly for the behavior of the tail regions that greatly influence estimation. An important outcome of this improvement is that an ideal observer analysis of sound location using responses from AI neurons can now employ parametric probability models using the proper dependence structure and marginals (Jenison, 2000; Jenison & Reale, 2003).

Shannon information can provide a measure of the information available for localizing a sound source in space transmitted on average by a given ensemble of $N$ neurons. Mutual information between stimulus and response is based on Shannon's entropy, which can be interpreted as the degree to which the total response entropy is reduced by the entropy conditioned on a particular stimulus setting $\theta$. Let this conditional entropy be $H[x_1, x_2, \ldots, x_N \mid \theta]$, where $x_i$ is the response of the $i$th neuron. Using the copula density, it is straightforward to show that the conditional entropy can be split into two terms: the sum of entropies due to the independent components of the conditional density and the quantity due strictly to the dependence structure defined by the copula (see appendix B). This can be expressed as

$$H[x_1, x_2, \ldots, x_N \mid \theta] = \sum_{i=1}^{N} H[x_i \mid \theta] + H_c[u_1, u_2, \ldots, u_N \mid \theta], \tag{2}$$

where the copula entropy is

$$H_c[u_1, u_2, \ldots, u_N \mid \theta] = -\int_{[0,1]^N} \mathrm{c}[u_1, u_2, \ldots, u_N \mid \theta] \log \mathrm{c}[u_1, u_2, \ldots, u_N \mid \theta] du. \tag{3}$$

It follows as a consequence of equation 2 that copula entropy must be mathematically equivalent to the negative of the mutual information between neurons, but with the benefit of being computed directly from the dependence structure. Figure 1C shows the reduction in the conditional response entropy due to the AMH copula as a function of dependence and parameterized by ensemble size $N$. The entropy of the copula is at its maximum at zero dependence, and the copula entropy declines as a function of dependence and ensemble size. So although the decline is rather modest for a pair of neurons, as suggested by Nirenberg et al. (2001), it accelerates here as the ensemble size increases.

The importance of this separability is that it is generally complicated, and often intractable, to parametrically model the joint probability density if it is nongaussian. When independence is assumed and the product density is based on only the marginal densities, the joint probability is easy to compute. Spike latency joint probability density can now be constructed using the marginal densities coupled by the separately estimated copula density function, which greatly simplifies the construction. The gaussian assumption is no longer a necessary constraint, and this newly found freedom allows for much greater flexibility in computational modeling studies of the brain. Finally, given the separability, the dependence entropy can be computed directly to assess the impact of the shape of neural dependence on sensory coding.

## Appendix A: Ali-Mikhail-Haq (AMH) Copula

The AMH copula distribution (Ali, Mikhail, & Haq, 1978) is an Archimedean copula (Nelsen, 1999) that can be constructed by an additive generator function $\varphi$ and its inverse $\varphi^{-1}$, if it exists, using the following form:

$$\mathbf{C}_{Archimedean}[u_1, u_2, \ldots, u_N; \alpha] = \varphi^{-1}\left[\sum_{i=1}^{N} \varphi[u_i; \alpha]\right]. \tag{A.1}$$

The generator function for the AMH copula is $\varphi[u; \alpha] = \log[\frac{1-\alpha(1-u)}{u}]$, which has an inverse $\varphi^{-1}[y; \alpha] = \frac{1-\alpha}{e^y - \alpha}$, and yields

$$\mathbf{C}_{AMH}[u_1, u_2, \ldots, u_N; \alpha] = \frac{\alpha - 1}{\alpha - \prod_{i=1}^{N} \frac{1-\alpha+\alpha u_i}{u_i}}, \tag{A.2}$$

where $u_i = F_i(x_i) = \int_{-\infty}^{x_1} p_i(t)\, dt$ is the marginal cumulative distribution. $\alpha$ is the measure of dependence where $0 \leq \alpha < 1$. An Archimedean $n$-dimensional copula is a proper distribution if and only if the inverse generator function $\varphi^{-1}$ is completely monotonic (Schweizer & Sklar, 1983).

## Appendix B: Copula Differential Entropy

The joint entropy can be factorized into the sum of entropies due to each marginal density and the entropy of the copula (dependence structure):

$$H[x_1, x_2, \ldots, x_N] = \sum_{i=1}^{N} H[x_i] + H_c\,[u_1, u_2, \ldots, u_N] \tag{B.1}$$

$$\begin{aligned}
&- \int_{\mathbb{R}^N} p[x_1, x_2, \ldots, x_N] \log[p[x_1, x_2, \ldots, x_N]]\, dx \\
&\quad = - \int_{\mathbb{R}^N} \prod_{i=1}^{N} p_i[x_i] \times c\,[u_1, u_2, \ldots, u_N] \\
&\qquad \times \log\left\{\prod_{i=1}^{N} p_i[x_i] \times c\,[u_1, u_2, \ldots, u_N]\right\} dx \\
&\quad = - \int_{\mathbb{R}^N} \prod_{i=1}^{N} p_i[x_i]\, c\,[u_1, u_2, \ldots, u_N] \\
&\qquad \times \left\{\sum_{i=1}^{N} \log p_i[x_i] + \log c\,[u_1, u_2, \ldots, u_N]\right\} dx.
\end{aligned} \tag{B.2}$$

Integrate both sides of the equation over the unit hypercube:

$$-\int_{[0,1]^N}\int_{\mathbb{R}^N} p[x_1, x_2, \ldots, x_N]\log[p[x_1, x_2, \ldots, x_N]]\,dx\,du$$

$$= -\int_{[0,1]^N}\int_{\mathbb{R}^N}\prod_{i=1}^{N} p_i[x_i]\,c\,[u_1, u_2, \ldots, u_N]$$

$$\times\left\{\sum_{i=1}^{N}\log p_i[x_i] + \log c\,[u_1, u_2, \ldots, u_N]\right\} dx\,du \tag{B.3}$$

$$-\int_{\mathbb{R}^N} p[x_1, x_2, \ldots, x_N]\log[p[x_1, x_2, \ldots, x_N]]\,dx$$

$$= -\sum_{i=1}^{N}\int p_i[x_i]\log p_i[x_i]\,dx_i$$

$$-\int_{[0,1]^N} c\,[u_1, u_2, \ldots, u_N]\log c[u_1, u_2, \ldots, u_N]\,du \tag{B.4}$$

$$H[x_1, x_2, \ldots, x_N] = \sum_{i=1}^{N} H[x_i] + H_c[u_1, u_2, \ldots, u_N] \tag{B.5}$$

## Acknowledgments

This work was supported by National Institutes of Health grant DC-03554.

## References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Auto. Control., 19*, 723.

Ali, M. M., Mikhail, N. N., & Haq, M. S. (1978). A class of bivariate distributions including the bivariate logistic. *J. Multivariate Anal., 8*, 405–412.

Brugge, J. F., Reale, R. A., & Hind, J. E. (1996). The structure of spatial receptive fields of neurons in primary auditory cortex of the cat. *J. Neurosci., 16*, 4420–4437.

Casella, G., & Berger, R. L. (1990). *Statistical inference*. North Scituate, MA: Duxbury Press.

Jenison, R. L. (2000). Correlated cortical populations can enhance sound localization performance. *J. Acoust. Soc. Am., 107*, 414–421.

Jenison, R. L., & Reale, R. A. (2003). Likelihood approaches to sensory coding in auditory cortex. *Network: Comp. Neural Sys., 14*, 83–102.

Jenison, R. L., Reale, R. A., Hind, J. E., & Brugge, J. F. (1998). Modeling of auditory spatial receptive fields with spherical approximation functions. *J. Neurophysiol., 80*, 2645–2656.

Joe, H. (1997). *Multivariate models and dependence concepts*. London: Chapman & Hall.

Mari, D. D., & Kotz, S. (2001). *Correlation and dependence*. London: Imperial College Press.

Nelsen, R. B. (1999). *An introduction to copulas*. New York: Springer-Verlag.

Nirenberg, S., Carcieri, S. M., Jacobs, A. L., & Latham, P. E. (2001). Retinal ganglion cells act largely as independent encoders. *Nature, 411*, 698–701.

Pola, G., Thiele, A., Hoffmann, K. P., & Panzeri, S. (2003). An exact method to quantify information transmitted to different mechanisms of correlational coding. *Network: Comp. Neural Sys., 14*, 35–60.

Reale, R. A., Jenison, R. L., & Brugge, J. F. (2003). Directional sensitivity of neurons in the primary auditory (AI) cortex: Effects of sound-source intensity level. *J. Neurophysiol., 89*, 1024–1038.

Schweizer, B., & Sklar, A. (1983). *Probabilistic metric spaces*. Amsterdam, North-Holland.

Sklar, A. (1959). Fonctions de répartition à $n$ dimensions et leur marges. *Publ. Int. Stat Univ., Paris, 8*, 229–231.

Soofi, E. S. (2000). Principal information theoretic approaches. *J. Am. Stat. Assoc., 95*, 1349–1353.

Yule, G. U. (1897). On the theory of correlation. *J. Roy. Statist. Soc., 60*, 812–854.

Received May 29, 2003; accepted September 16, 2003.